The Importance of Functional Form in the Estimation of Welfare

Catherine L. Kling

Researchers have recognized the central role that the choice of functional form has on estimates of consumer surplus. The purpose of this paper is to quantify the magnitude of errors which might arise from the use of incorrect functional forms. It describes a simulation experiment where estimated consumer surplus, based on simulated data sets, is compared with consumer surplus computed directly from the simulated data. The errors resulting from the use of mismatching functional forms range from approximately 4% to 107%.

Key words: benefit estimation, travel cost models, consumer surplus, functional form, recreation demand.

For over two decades applied economists have been estimating environmental benefits using recreation demand models. The approaches used to estimate benefits have evolved from relatively simple, single-equation models of demand to more complex models incorporating, among other things, multiple sites, truncation problems, the opportunity cost of time, and environmental quality variables. Underlying this evolution has been a shift in emphasis from ad hoc specifications of the demand functions to demand functions which are consistent with the postulates of consumer theory.

This paper explores the importance of one component of these models for the estimation of benefits: namely, the choice of functional form. Researchers have recognized the central role that the choice of functional form has on benefit estimates. However, the magnitude of error associated with incorrect functional form has not been measured. This paper presents a first attempt at quantifying these errors.

To accomplish this objective, a simulation experiment is conducted wherein three sets of individual observations on recreation users are generated based on linear, semilog, and double-log demand functions, respectively. Consumer surplus (*cs*) associated with the current quantity of use is calculated. These simulated data sets are then treated as actual data sets, and the three demand functions are estimated on each of the three data sets. Estimates of *cs* based on these estimated demand functions are compared to the simulated *cs* measures calculated from the simulated data. In this way estimates of *cs* can be directly compared to the "true" *cs*, and the errors resulting from the use of unmatching or "incorrect" functional forms can be quantified. The procedure is replicated 50 times to examine the robustness of the results.[1]

Following a discussion of the current approaches to specifying functional form in the first section, the results of the simulation experiment are presented in the second section. To test the sensitivity of the simulation results to the size of the welfare change, the third section contains the results of the simulation experiment when small price changes are considered. Preliminary findings and conclusions are discussed in the final section.

The author is an assistant professor of agricultural economics at the University of California, Davis.

This is Giannini Foundation Research Paper No. 881.

The helpful comments and insights of Richard Sexton and an anonymous reviewer are gratefully acknowledged.

[1] Additional discussion of the simulation approach is contained in Kling 1988a and Kling 1988b.

The Choice of Functional Form and Welfare Measurement

The empirical importance of functional form has been noted by many authors (for example, Bockstael and McConnell; Ziemer, Musser, and Hill; Sutherland). Ziemer, Musser, and Hill note that different functional forms can generate very different magnitudes of benefits. They compare *cs* estimates from linear, semilog, and quadratic demand functions. In their results, these different functional forms yield average *cs* estimates that range from \$20.46 (linear) to \$79.09 (semilog).

Examination of the recreation demand literature suggests that researchers have employed at least three different ways to choose functional forms. First, the researcher may simply choose a form, based on ease of estimation, intuition, or previous knowledge of the data. Demand is then estimated using this form, and welfare estimates are produced. This is the simplest and perhaps most common method of selecting a demand function.

A second approach is to estimate several functional forms. Goodness-of-fit tests are then performed to distinguish among the estimated forms. The best-fitting function is assumed to be preferable for welfare estimation. Given the difficulties associated with goodness-of-fit tests, researchers typically choose forms that can be nested easily or which can be compared based on the Box-Cox test. This approach is consistent with the long-held premise that researchers should "let the data tell the story."

This approach also appears to be the one Hausman espouses in his influential paper concerning the recovery of Hicksian measures of welfare from Marshallian demand functions:

> First, the only observable data are the market demand data so good econometric practice would indicate finding a function that fits the data well. Thus, different specifications of the demand curve, not the utility function, would be fit with the best-fitting demand equation chosen to base the applied welfare analysis on . . . (page 664)

Goodness-of-fit tests done to date generally support the use of a semilog functional form. Ziemer, Musser, and Hill employ conventional hypothesis testing procedures to distinguish between the linear and quadratic. These authors cannot reject the hypothesis that there is a significant difference between these two forms. Employing a Box-Cox transformation, they conclude that the semilog better fits their data than the linear specification.

In addition to the work by Ziemer, Musser, and Hill, several other authors have employed goodness-of-fit tests to empirically compare different forms. In general, these studies corroborate Ziemer, Musser, and Hill's results supporting the choice of the semilog functional form. Using nonnested hypothesis testing procedures, Smith finds that a semilog best fits his data. Strong and Vaughan, Russell, and Hazilla also select the semilog, noting that questions of functional form and heteroskedasticity are interrelated. McConnell concludes that to date the semilog is the form most supported by empirical evidence.

A dissenting view is reported by Adamowicz, Fletcher, and Graham-Tomasi. Based on considerations of the variance of the welfare measure, these authors conclude that the double log yields *cs* estimates which are less sensitive to changes in the travel cost parameter than other forms.

The authors discussed above have been concerned primarily with choosing from among common, ad hoc demand specifications such as the linear, semilog, double log, or quadratic. In contrast, a third route to determining functional form is to start from a utility function. In this approach, the researcher chooses a utility function which he or she believes contains desirable properties. Demand functions then are derived from this utility function and estimated. This approach guarantees that the resulting demand functions will satisfy the integrability conditions; i.e., that the Slutsky matrix is symmetric and negative semidefinite. This approach appears to have gained support recently; authors such as Morey and Kealy and Bishop provide examples.

The criterion for determining the "best" functional form differs considerably between the latter two approaches. A researcher subscribing to the goodness-of-fit approach would choose the functional form that provides the closest fit of the data; the researcher following the utility theoretic approach would choose the form that is generated by the most sensible utility function. These two approaches are not necessarily mutually exclusive; it is possible for a researcher to estimate several utility theoretic demand functions, perform goodness-of-fit tests, and choose the form that best fits the data. However, this approach generally has not been applied.

The work of Ziemer, Musser, and Hill emphasizes the practical significance of these two approaches. If welfare measurement were not highly dependent upon functional form, whether one took a utility theoretic approach or a goodness-of-fit approach only would be of consequence in the ivory tower. However, welfare measurement appears to be highly dependent on functional form.

Which approach to choice of functional form is preferable is both a theoretical and empirical question. In the simulation experiments presented here, the errors associated with incorrect specification of the demand function are examined. This exercise does not address directly the question of which approach, the utility theoretic or demand based, is to be preferred; however, the results of the simulations presented below may provide some useful insights into the resolution of this debate.

Errors in Welfare Measures Associated with Incorrect Functional Forms

To measure the errors in welfare estimates resulting from the use of incorrect functional forms, a simulation experiment is conducted. A simulated data set is generated by choosing a demand function, parameter values, and an error structure. That information is then combined with price and income data to yield simulated quantities of recreational visits. Measures of *cs* associated with a $60 price increase are calculated.

A fixed price increase is employed rather than the entire area under the demand curve for consistency among different demand functions. Since the area under some demand functions is unbounded, it is necessary to choose an upper price limit under which to integrate. The use of $60 as this limit in all cases assures comparability of the results. A complete simulated data set is composed of quantity, price, income, and *cs* information.

Two hundred simulated observations are generated by this process for each data set. Three different demand functions—linear, semilog, and double log—are employed in separate experiments to generate data sets.

The second stage of the simulation experiment entails using the simulated data to estimate the three demand functions and calculate estimates of *cs* based on each. In each of the simulations, one demand function is an exact match of the function used to generate the simulated data. For example, a linear demand function is used to generate a data set upon which linear, semilog, and double-log demand functions are estimated. Consumer surplus is calculated for both the simulated data and the three estimated demand functions. The simulated and estimated surplus measures are then compared. This exercise is repeated using both the semilog and double-log demand functions to generate the data. In this way, the importance of an exact match of the functional form for welfare estimation can be quantified.

To perform Monte Carlo repetitions, an error term is imbedded in the demand function and 50 repetitions of each experiment are performed. Since the size and structure of the error term introduced into the simulated data may be critical to the outcome, three different assumptions concerning the error structure are employed. In total, nine sets of repetitions are performed: three demand functions with three different error structures are combined to generate the complete experiment.

The forms of the demand functions are

(1) the linear:
$$x_i = \alpha_1 + \beta_1 p_i + \gamma_1 y_i,$$
(2) the semilog:
$$\ln(x_i) = \alpha_2 + \beta_2 p_i + \gamma_2 y_i,$$
(3) the double log:
$$\ln(x_i) = \alpha_3 + \beta_3 \ln(p_i) + \gamma_3 \ln(y_i),$$

where x_i is the number of recreation visits taken by simulated individual i, p_i is the price associated with a trip, y_i is income, and the Greek letters correspond to parameters. Two hundred simulated observations are generated for each data set.

The *cs* associated with a price increase of $60 is calculated for each simulated observation. The *cs* formulas are

(4) the linear:
$$cs_1 = \frac{1}{2\beta_1}(x1_i^2 - x0_i^2),$$
(5) the semilog:
$$cs_s = \frac{x1 - x0}{\beta_2},$$
(6) the double log:
$$cs_d = \frac{p_1 x1 - p_0 x0}{\beta + 1},$$

Table 1. Functional Forms, Error Structures, and Coefficient Values Employed in the Simulation Experiments

Functional Form	Error Structure	Coefficient Values α	β	γ
Linear	additive: $\epsilon_i \sim N(0,1)$	11	$-.25$	.0002
	$\epsilon_i \sim N(0,9)$	11	$-.25$	.0002
	mult: $\epsilon_i \sim N(0,.01)$	20	$-.43 + \epsilon_i$	.00006
Semilog	additive: $\epsilon_i \sim N(0,.111)$	2	$-.05$	.00001
	$\epsilon_i \sim N(0,1)$	2	$-.05$	.00001
	mult: $\epsilon_i \sim N(0,.009)$	2	$-.12 + \epsilon_i$	.00003
Double log	additive: $\epsilon_i \sim N(0,.111)$	1	$-.75$	.3
	$\epsilon_i \sim N(0,1)$	1	$-.75$	.3
	mult: $\epsilon_i \sim N(0,.04)$	1	$-.75 + \epsilon_i$	.3

where p_0, p_1 are the prices before and after the price increase, and $x0$ and $x1$ are the number of recreational visits before and after the price increase, respectively.

Fifty repetitions of each data set are generated by employing 50 different sets of random errors based on the same distribution. For each demand function, a relatively large additive normal error, a relatively small additive normal error, and a multiplicative error are used. The additive error enters the demand as follows: $x_i = f(p_i, y_i) + \epsilon_i$.

In contrast to the introduction of the additive error, in a third set of repetitions the error is directly added to the price coefficient; that is, $(\beta + \epsilon_i)$. Table 1 lists the functional forms, error structures, and coefficient values employed in the simulations.

The data on prices and income used in the simulation experiment come from Chesapeake Bay data on recreational beach use in the summer of 1984 compiled by Research Triangle Institute for the University of Maryland.[2] The incomes range from \$7,500 to \$120,000, and the average is \$41,287. The prices range from \$2 to \$25.10 and average \$15.

The Marshallian measure of *cs* is employed rather than either compensating or equivalent variation for several reasons. First, for some functional forms, no general closed-form solution for compensating or equivalent variation exists. Second, drawing from Willig's pioneering work, calculations of the differences between the Marshallian measure and the Hicksian measure are, in most cases, a fraction of 1%. These errors are dwarfed by the errors generated as a result of incorrect functional form.[3] Finally, Marshallian measures commonly are employed in the recreation demand literature. Since the Willig bounds indicate trivial differences between the measures, the Marshallian measure is employed.

Table 2 contains the results of comparisons between the simulated consumer surpluses and the estimated consumer surpluses for the three linear cases. The first box contains the results of the simulation employing the linear demand function and a small additive error. Using the linear simulated data, linear, semilog, and double-log demand functions are estimated and the resulting estimates of *cs* are compared with the simulated *cs*.

The first column in table 2 contains the average over the 50 repetitions of the *cs* estimates. The second column measures the average difference (the mean error) between the simulated *cs* and the estimated *cs*. In this case, the linear demand function results in a mean error (*ME*) of only $-$\$1.24. The semilog results in an overestimate, on average, of *cs* of \$329.43, and the double log results in an overestimate of \$267.39.

The last column in table 2 contains the mean error as a percentage of the simulated consumer surplus (*MEC*). The *MEC* is calculated to facilitate comparisons across the different simulations since it expresses the errors in percentage terms. For the linear additive error simulation, the linear, semilog, and double log generate approximately $-.25\%$, 66% and 54% errors, respectively.

[2] For a complete description of the data, see Bockstael, Hanemann, and Strand.

[3] Calculations of the Willig bounds for the nine sets of repetitions described here indicate that the differences between the Hicksian measures and the Marshallian measures are less than .31% in all cases and in most cases range from .01% to .20%.

Table 2. Summary Statistics for the Simulation Experiments

Functional Form[a]	*CSP*	*ME*	*MPE*	*RMSE*	*RMSPE*	*MEC*
Linear Simulated Demand Function (small additive error)						
Linear	$495.74	−1.24	−0.00	12.53	0.03	−0.25
Semilog	826.41	329.43	0.95	330.62	1.12	66.29
Double log	764.37	267.39	0.83	272.22	0.03	53.80
Linear Simulated Demand Function (large additive error)						
Linear	$497.19	−8.98	−0.02	38.17	0.09	−1.77
Semilog	769.77	263.60	0.95	272.12	1.46	52.08
Double log	751.84	245.56	0.95	259.69	1.52	48.51
Linear Simulated Demand Function (multiplicative error)						
Linear	$316.78	−25.11	0.02	102.20	0.23	−7.34
Semilog	584.25	242.33	1.12	263.12	1.44	70.88
Double log	716.36	374.47	1.68	390.82	2.11	109.53

[a] CSP = mean predicted $cs = (1/N)\,\Sigma_i\, csp_i$, $i = 1, \ldots, N$; CS = mean simulated surplus = $(1/N)\,\Sigma_i\, cs_i$; ME = mean error = $CSP - CS$; MPE = mean proportional error = $(1/N)\,\Sigma_i(csp_i - cs_i)/cs_i$; $RMSE$ = root-mean-squared-error = $[(1/N)\,\Sigma(csp_i - cs_i)^2]^{1/2}$; $RMSPE$ = root-mean-squared-proportional-error = $[(1/N)\,\Sigma((csp_i - cs_i)/cs_i)^2]^{1/2}$; and $MEC = ME/CS$, where csp_i and cs_i are the predicted and simulated consumer surplus for individual i, and N is the number of observation in all repetitions (10,000).

Since it is the total estimates of benefits which are generally of interest, most attention will be spent discussing the statistics measuring the average error. However, for some applications the variance of the estimate may be of importance (Adamowicz, Fletcher, and Graham-Tomasi). Therefore, measures of the root-mean-square-error (*RMSE*) and root-mean-square-percentage-error (*RMSPE*) also are presented.

The large additive error and multiplicative error specifications yield similar results. The *MEC*s in both cases are much larger when the semilog and double-log demand functions are employed. The estimates of the latter two forms yield errors ranging from about 52% to over 109% of the simulated *cs*.

Table 3. *MECs* for the Linear, Semilog, and Double-log Simulations for the Entire Consumer Surplus

Functional Form	Linear	Semilog	Double log
Linear			
Small add. error	−0.25	66.29	53.80
Large add. error	−1.77	52.08	48.51
Multiplicative	−7.34	70.88	109.53
Semilog			
Small add. error	−30.60	5.34	81.17
Large add. error	26.94	15.01	85.42
Multiplicative	52.94	−14.49	68.91
Double log			
Small add. error	−63.49	−42.70	−0.00
Large add. error	−32.40	−43.25	−2.80
Multiplicative	−63.79	−51.09	−12.63

To facilitate comparison with other functional forms, table 3 contains the *MEC*s from the three linear simulations. To save space, only the *MEC*s for the semilog and double-log simulations are presented.

The semilog simulations exhibit a slightly different pattern than the linear simulations. In the small additive error case, the matching functional form results in the smallest *MEC*, with 5.34. This is considerably larger than the almost zero *MEC* in the small additive case for the linear demand. Further, in the large additive error and multiplicative error cases, the *MEC*s are about 15%. Once again, the mismatching functional forms yield large *MEC*s ranging in absolute value from 27 to 85.[4]

The double-log simulations display a similar pattern. In each of the three cases, the matching functional form results in the smallest *MEC*s and the mismatching forms result in much larger percentage errors. The *MEC*s of the mismatching forms range in absolute value from about 32 to 64.

Overall, the results from table 3 indicate that the error from employing a mismatching functional form can be quite large, often exceeding 40% and reaching a high of 109%. However,

[4] In the version of this paper presented at the 1988 WAEA meetings in Honolulu, the semilog *cs* estimates were calculated using the predicted number of visits rather than the actual. Since the simulated *cs* is calculated using the actual number of visits, the use of the predicted number resulted in a few large errors and generated some anomalies in the results. In particular, in two cases in the semilog simulation, the matching functional form performed worse than the unmatching forms.

Table 4. Estimated Demand Functions for the First Repetition

Simulated Model	Estimated Model	Intercept	P (or ln$P$)	y (or ln$y$)	R^2
Linear small error	Linear	11.29	−0.26	.00020	.96
	Semilog	2.48	−0.02	.00001	.91
	Double log	−2.18	−0.18	.51	.92
Linear large error	Linear	10.25	−0.21	.00019	.73
	Semilog	2.41	−0.02	.00001	.58
	Double log	−2.43	−0.18	.53	.58
Linear mult. error	Linear	9.53	−0.40	.00061	.99
	Semilog	2.69	−0.02	.00002	.88
	Double log	−5.66	−0.19	.89	.95
Semilog small error	Linear	7.25	−0.30	.000085	.54
	Semilog	1.96	−0.05	.000011	.62
	Double log	−1.56	−0.48	.43	.55
Semilog large error	Linear	12.47	−0.39	.00005*	.06
	Semilog	1.99	−0.04	.000006	.06
	Double log	0.05*	−0.42	.26	.07
Semilog mult. error	Linear	6.72	−0.81	.00032	.59
	Semilog	1.89	−0.13	.000033	.86
	Double log	−10.10	−1.24	1.40	.80
Double-log small error	Linear	21.43	−0.88	.000058	.57
	Semilog	2.98	−0.07	.000007	.64
	Double log	.70	−0.76	.33	.66
Double-log large error	Linear	62.67	−3.05	.000079*	.15
	Semilog	3.42	−0.10	.000004*	.25
	Double log	2.02*	−1.01	.26	.27
Double-log mult. error	Linear	20.43	−0.73	.000047	.33
	Semilog	2.89	−0.06	.000006	.35
	Double log	.65*	−0.68	.31	.39

Note: the coefficients indicated with an asterisk are not significant at the .05 level—all other coefficients are significant at this level.

even when matching forms are used, there is no guarantee of a perfect fit of the welfare estimate; matching forms result in *MEC*s ranging in absolute value from zero to 15.

To give an idea of how goodness of fit of the estimated demand equations affects the *MEC*s, table 4 contains the results of estimates of the three demand functions for the first repetition of each experiment. The demand functions estimated in the remaining 49 repetitions are very similar to those reported in table 4.

Small Price Changes

To examine the accuracy of welfare measures when the price changes are small rather than large, the experiments are repeated. The experiments are identical in all respects except that a $5 price change is employed rather than a $60 price change. In this way, it is possible to determine the importance of functional form choice when the goal is welfare measurement for a small price change.[5]

The results indicate in general that for a small price change, the choice of functional form is not nearly as critical as it is for a large change. In most cases, the errors from employing unmatching functional forms are less than a few percentage points.

For example, when the linear demand function is employed to generate the data, the average *MEC*s for the linear-estimated model are all approximately zero. Further, the *MEC*s from the semilog and double-log models range from .04 to 2.71 in absolute value. In other words, the largest average error generated by using an incorrect functional form is only 2.71%.

[5] Actually, the $5 price change employed here is not all that small. Since the average original price in the data is roughly $15, a $5 price increase corresponds to a 33% price jump. It appears that the error in welfare estimates increases at an increasing rate with price.

The errors are not quite as small in the semilog and double-log based simulations. In the semilog simulations, the average *MEC*s from the mismatching forms range from .61 to 38.16 in absolute value with an overall mean of 13.13. Likewise in the double-log simulations, the average *MEC*s range from 2.48 to 39.11 in absolute value with a mean of 11.40.

It appears that the choice of functional form is less critical for welfare measurement purposes in the case of a small price change.

Conclusions and Discussion

This paper presents evidence on the importance of functional form in the estimation of welfare. In the simulations presented here, the use of unmatching functional forms results in errors to welfare measures ranging from 26.94% to 109.53%. These results dramatize the very large effect the choice of functional form can have on welfare estimates.

The examination of errors from unmatching functional forms for small price changes results in a different picture. For a $5 price change, the *MEC*s averaged over all mismatching forms is only 8.53. These results suggest that, as one would expect, the choice of functional form is less critical when a smaller price change is proposed.

The results presented here are based on particular demand functions, parameter values, and error structures and are subject to the qualification that they may not be generalizable to all cases. However, they do point out the potential size of errors that may arise from misspecification. Further, they are based on three commonly employed functional forms and three different assumptions concerning the introduction of the error term.

In practice, researchers employing goodness-of-fit tests to choose among functional forms may reduce the potential for large errors so that even if large price changes are desired, the researcher may feel confident in the estimate. The results presented here suggest that it well may be worth researchers' time to do goodness-of-fit tests when welfare evaluation is the goal of estimation. Additionally, the use of flexible functional forms, particularly those which are globally flexible, may prove useful. The use of goodness-of-fit tests and/or flexible functional forms to improve the reliability of welfare estimates are empirical questions which could be explored in the simulation context.

[*Received July 1988; final revision received September 1988.*]

References

Adamowicz, W., J. Fletcher, and T. Graham-Tomasi. "Functional Form and the Statistical Properties of Welfare Measures." Dep. Agr. Econ., Working Paper #87-28, Purdue University, March 1988.

Bockstael, N. E., and K. E. McConnell. "Calculating Equivalent and Compensating Variation for Natural Resource Environments." *Land Econ.* 56(1980):56–63.

Bockstael, N. E., W. M. Hanemann, and I. E. Strand. *Measuring the Benefits of Water Quality Improvements Using Recreation Demand Models.* Report presented to the Environmental Protection Agency under Cooperative Agreement CR-81143-01-1, Volume II, 1986.

Hausman, J. "Exact Consumer's Surplus and Deadweight Loss." *Amer. Econ. Rev.* 71(September 1981):662–76.

Kealy, M. J., and R. Bishop. "Theoretical and Empirical Specification Issues in Travel Cost Demand Studies." *Amer. J. Agr. Econ.* 68(1986):660–67.

Kling, C. L. "Comparing Welfare Estimates of Environmental Quality Changes from Recreation Demand Models." *J. Environ. Econ. and Manage.* 15(1988a):331–40.

———. "The Reliability of Estimates of Environmental Benefits from Recreation Demand Models." *Amer. J. Agr. Econ.* 70(1988b):892–901.

McConnell, K. E. "The Economics of Outdoor Recreation." *Handbook of Natural Resource and Energy Economics,* eds., Allen Kneese and James Sweeney. New York: Elsevier Science Publishers, 1985.

Morey, E. "The Choice of Ski Areas: Estimation of a Generalized CES Preference Ordering with Characteristics." *Rev. Econ. and Statist.* 66(1984):584–90.

Smith, V. K. "Travel Cost Demand Models for Wilderness Recreation: A Problem of Non-Nested Hypotheses." *Land Econ.* 51(1975):103–11.

Strong, E. "A Note on the Functional Form of Travel Cost Models with Zones of Unequal Populations." *Land Econ.* 59(1983):342–49.

Sutherland, R. J. "The Sensitivity of Travel Cost Estimates of Recreation Demand to the Functional Form and Definition of Origin Zones." *West. J. Agr. Econ.* 7(1982):87–98.

Vaughan, W., C. Russell, and M. Hazilla. "A Note on the Use of Travel Cost Models With Unequal Zonal Populations: Comment." *Land Econ.* 58(1982):400–07.

Willig, R. "Consumer's Surplus Without Apology." *Amer. Econ. Rev.* 66(September 1976):589–97.

Ziemer, R., W. Musser, and R. C. Hill. "Recreation Demand Equations: Functional Form and Consumer Surplus." *Amer. J. Agr. Econ.* 62(1980):136–41.